Gear's algorithm

As long as we're talking about matrix methods, let's also look at Gear's very clever algorithm for segmenting sets of points with independent motions.

First, we need to know that if we have a set of points undergoing 3D motion, scaling and orthographic projection, we can write this as:

$I = SP$

Where I is a 2xn matrix that gives the image coordinates of n points, S is a 2x4 matrix that encodes rotation, translation and scaling, and P is a 4xn matrix in which each column contains the x-y-z coordinates of a 3D point, with a row of all 1's at the bottom.

If we have k frames, we can make S a (2k)xn matrix, in which pairs of rows represent the motion in different frames. Then I becomes a (2kxn) matrix, in which each column represents the x-y coordinates of a point over all frames.

It is important that I has only rank 4, even though it is big. This is because the world is 3-Dimensional.

Now, suppose we have some measurement matrix W, which comes from concatenating together the I matrices we get from M independently moving objects. In general, the columns associated with one object will be linearly independent of the columns from another object, so W will have rank 4M. We would like to figure out which columns belong together, that is, which columns have a shared motion.

We explain this idea in the absence of sensor noise, and when all points are in general position. We go column-by-column, selecting a basis for the column-space of W. In general, the first four columns will be linearly independent. The fifth column will be independent of the first four, unless this point shares a motion with all the other four. If it is independent of the first four columns, we add it to our basis. Eventually, we will get a column that shares a motion with four previously chosen columns. This column will be linearly dependent on these previous four. This immediately tells us we have a group of five points that share a motion. Eventually, we wind up with four columns in our basis for every motion, and every other column expressed as a linear combination of four prior ones. This makes the grouping of columns obvious; the linearly dependent ones are grouped with the four basis columns.

Things are a little more complicated when we have noise. Nothing is exactly linearly dependent on anything else, and a column may by chance be nearly dependent on four random columns. Gear solves this by using linear algebra methods to select the columns in the most stable way, and combining evidence we get from other columns being almost expressible as linear combinations of four basis elements. Since Gear's paper for details.